

H₂ emission as a diagnostic of physical processes in starforming galaxies

By Paul P. VAN DER WERF¹

¹Leiden Observatory, P.O. Box 9513, NL - 2300 RA Leiden, The Netherlands

Observations and interpretation of extragalactic rotational and rovibrational H₂ emission are reviewed. Direct observations of H₂ lines do not trace bulk H₂ mass, but excitation rate. As such, the H₂ lines are unique diagnostics, if the excitation mechanism can be determined, which generally requires high-quality spectroscopy and suitable additional data. The diagnostic power of the H₂ lines is illustrated by two cases studies: H₂ purely rotational line emission from the disk of the nearby spiral galaxy NGC 891 and high resolution imaging and spectroscopy of H₂ vibrational line emission from the luminous merger NGC 6240.

1. Introduction

Direct observations of H₂ emission from external galaxies have become standard practice in the past decade through the revolution in ground-based near-infrared instrumentation. As a result, the near-infrared H₂ rovibrational lines are now readily detectable throughout the local universe (e.g., Moorwood & Oliva 1988, 1990; Puxley *et al.* 1988, 1990; Goldader *et al.* 1995, 1997; ? (?). More recently, the Short Wavelength Spectrograph (SWS) on the Infrared Space Observatory (ISO) has for the first time allowed detection of the purely rotational H₂ lines in the mid-infrared spectral regime. For instance, the first detection (outside the solar system) of the H₂ S(0) line at 28.21 μm was reported by ? (?) from the star forming nucleus of the nearby spiral galaxy NGC 6946.

The interpretation of these data is, however, far from trivial. At typical molecular cloud temperatures ($T \sim 20$ K) the upper levels of even the lowest H₂ transitions are essentially unpopulated, and hence H₂ emission is, unlike for instance CO $J = 1 \rightarrow 0$ emission, not a tracer of bulk molecular gas mass. Instead, an excitation mechanism capable of populating these energy levels is required. Furthermore, the excitation rate needs to be high enough to maintain an excited level population sufficient for producing detectable emission. Hence the H₂ line luminosities measure the *rate of excitation*; as such they provide unique and highly diagnostic information, that cannot be obtained in any other way. In addition, since the H₂ lines are forbidden quadrupole transitions with very small Einstein A values, the lines are optically thin; hence there is, in contrast to the situation with most other molecular lines, no need to solve a complicated, geometry-dependent radiative transfer problem for a physical interpretation of the H₂ lines. The near-infrared H₂ lines will of course suffer from extinction by dust, but this effect can usually be quantified adequately by using suitable ratios of hydrogen recombination lines, [Fe II] lines or H₂ rovibrational lines in the same spectral range, with accurately known intrinsic flux ratios.

The situation is complicated, however, by the fact that H₂ excitation can be brought about by a variety of mechanisms, which may all play a role. For instance, in a starburst galaxy, H₂ emission may be generated by UV-pumping in starforming regions or by shocks due to supernova remnants or outflows; moreover, these processes are expected to occur

to appear in “*Molecular hydrogen in space*”, eds. F. Combes & G. Pineau des Forêts, Cambridge University Press

together in the same small volume occupied by the starburst, and thus a combination of excitation mechanisms may be expected at every position. Generally, this combination will be difficult to separate (e.g., NGC 253 observations: ? (?; ? (?). In addition, shocks due to large-scale streaming motions in spiral arms, bars (favoured for the barred Seyfert 2 NGC 1068 by ? (? or merger-driven flows (e.g., in NGC 6240, ? (?; see also § 3) may play a role. Furthermore, X-ray excitation may be produced by multiple supernova remnants, or, if present, by an active galactic nucleus (proposed for e.g., Cyg A by ? (? and for Cen A by ? (? or a cooling flow (?). In the absence of sufficient spatial resolution for separating the various excitation mechanisms, spectral diagnostics must be used. Models for UV-pumped, shocked or X-ray excited H_2 emission have reached considerable predictive power; however, the densities of the emitting regions are often sufficiently high to thermalize the relevant level populations, quenching the typical signatures of non-thermal excitation processes. The only accessible diagnostics are then fluxes of lines with very high critical densities. However, these lines are intrinsically faint, and accurate fluxes for such lines require long integrations with large telescopes. The complexity of the problem is well illustrated by the H_2 emission of the merging system NGC 6240, which has been attributed to X-ray excitation (?), UV-pumping (?), shocks (?) and formation pumping (?). However, all of the non-thermal excitation processes relied on poorly measured fluxes of faint lines. More recent spectroscopy of NGC 6240 by ? (?) showed that only shock excitation can account for the H_2 rovibrational spectrum (see also § 3).

Generally a combination of accurate multi-line spectroscopy and suitable additional information such as high resolution spatial information is needed for a proper analysis of the dominant excitation mechanism and hence a physical analysis of the H_2 emission. A complicating factor when combining H_2 rovibrational and purely rotational lines is the fact that the dominant excitation mechanisms of these lines may be different, because of the different excitation requirements of these transitions: while the lowest rotational lines require $T > 100$ K and are excited at moderate densities ($n_{H_2} > 10^2 \text{ cm}^{-3}$), the rovibrational lines require $T > 2000$ K and $n_{H_2} > 10^4 \text{ cm}^{-3}$. The following sections will therefore present two case studies. First an analysis of the extended purely rotational H_2 emission from the nearby edge-on spiral galaxy NGC 891 is given (§ 2). Then, the more extreme conditions in the merger NGC 6240 are discussed, using new high resolution near-infrared data. The general implications of these two cases are discussed in § 4.

2. Extended H_2 rotational line emission in the disk of NGC 891

? (?) have observed seven positions in the disk of the nearby (distance $D = 9.5$ Mpc) edge-on spiral galaxy NGC 891 with the ISO SWS. The positions observed include the nucleus and positions spaced along the disk to galactocentric distances R of 8 kpc south of the nucleus and 11 kpc north of the nucleus. At the 11 kpc north position, the ^{12}CO $J = 1 \rightarrow 0$ (?; ?; ?), and dust emission (?) are barely detected. With ISO, the H_2 S(0) ($28.21 \mu\text{m}$) and S(1) ($17.03 \mu\text{m}$) lines were detected at all positions. These are the first detections of these lines outside starburst or active nuclei, with the exception of the detection in the disk of NGC 6946 reported by ? (?).

The simplest analysis of these data would assume that the emission is fully thermalized (i.e., the high-density limit is assumed) and arises from an isothermal gas layer with an ortho-para ratio of three, in agreement with the statistical weights of the ortho- and para-varieties. Under these assumptions the S(1)/S(0) ratio yields the temperature of the emitting region, which can be combined with the observed fluxes to give the column density of emitting H_2 , averaged over the SWS aperture. However, these simple

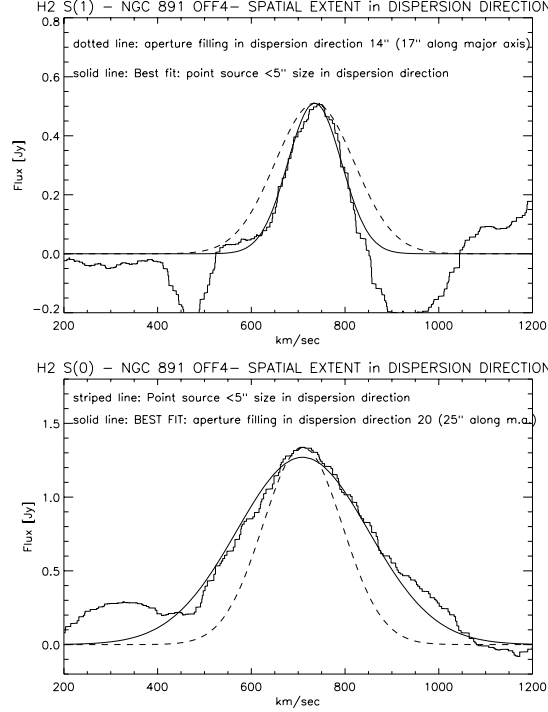


FIGURE 1. H_2 S(1) (upper panel) and S(0) (lower panel) lines from $R = 8$ kpc south in the disk of NGC 891, with fits for a point source and an aperture-filling source as indicated.

assumptions lead to unacceptable results. For instance, at the position 8 kpc north (a typical “disk” position), the S(0) and S(1) data then imply $N(H_2) = 2.7 \cdot 10^{22}$ at $T = 76$ K. This result is incompatible with the CO $J = 2 \rightarrow 1 / J = 1 \rightarrow 0$ line ratio of 0.75, at this position (?), which strongly rules out a large mass of dense molecular gas at $T = 76$ K. This argument shows that the simple assumptions used above do not suffice.

The discrepancy can be removed by allowing lower densities, and therefore subthermal excitation of the H_2 $J=3$ level, giving rise to a higher implied temperature for a given S(1)/S(0) ratio. Because the temperatures involved are still much lower than the upper levels of the transitions involved, the implied H_2 mass will be extremely sensitive to the adopted temperature, and thus strongly decrease as lower densities are allowed. An additional advantage of this procedure is that the CO $J=2$ level will now also be subthermally populated, so that CO $J = 2 \rightarrow 1 / J = 1 \rightarrow 0$ line ratios lower than the thermal value can be tolerated in the analysis. Indeed the multi-level multi-isotope CO data by ? (?) indicate dominant densities of $n_{H_2} \sim 10^3 \text{ cm}^{-3}$ in the nuclear region and $n_{H_2} \sim 200 \text{ cm}^{-3}$ throughout the disk of NGC 891. Adopting these densities, the ISO data imply H_2 column densities varying from 10 to 30% of the value derived from CO, at temperatures of 120 to 130 K, throughout the inner disk ($R < 8$ kpc) of NGC 891.

At the outer disk positions ($R = 8 - 11$ kpc) however, this solution is not adequate, since the S(0) and S(1) line profiles at these positions are significantly different. The SWS uses an aperture much larger than the diffraction beam and hence the observed line width depends on the extent of the emission region; the S(0) line shows the broad profile characteristic of aperture-filling emission, while the S(1) profile is narrow, indicating emission from a region much smaller than the aperture (Fig. 1). Thus the S(0) and S(1) lines in the outer disk of NGC 891 must originate in separate regions: a warm

component ($T > 130$ K) dominating the S(1) emission and located in isolated regions in the disk and a separate more pervasive component, dominating the S(0) line. The latter component may contain very significant mass, depending on its temperature, and the implications of this possibility for the mass budget at the outer positions have been explored by ? (?). However, a solution where the component dominating the S(0) line is cool ($T \sim 90$ K) is problematic, since the heating required for maintaining a very large mass of H_2 at $T \sim 90$ K cannot be provided. In thermal equilibrium the heating rate should equal the cooling rate, which is dominated by [C II] 158 μm emission (which at $T \sim 90$ K is an extremely efficient cooler, since the upper level of the 158 μm transition is at 91 K), with possible contributions from H_2 rotational lines, CO rotational lines, and [C I] emission. The [C II] 158 μm emission from NGC 891 has been observed by ? (?), and it is easily verified that in the outer disk the [C II] emission dominates the H_2 emission (SWS data discussed here), CO emission (barely detected in the outer disk) and [C I] emission (?). If all available carbon (assumed to have solar abundance) is in the form of C^+ , the [C II] emission allows a column density of at most $N(H_2) \sim 7 \cdot 10^{21} \text{ cm}^{-2}$ at $T \sim 90$ K in the outer disk, averaged over the SWS aperture, while $N(H_2) \sim 10^{23} \text{ cm}^{-2}$ would be required to produce the observed S(0) line flux.

This problem can be solved by relaxing the final assumption: that of an ortho-para of three. A lower ortho-para ratio raises the implied temperature and lowers the implied mass in the same way as a lower density does (as discussed above). Assuming an ortho-para ratio of unity, the implied mass is lowered sufficiently that the required heating can be accounted for.

This analysis thus favours an interpretation in which the H_2 rotational emission arises in low-density ($n_{H_2} \sim 200 \text{ cm}^{-3}$), warm ($T \sim 120$ K) molecular gas with an ortho-para ratio of about unity, which pervades the disk of NGC 891 at least to the end of the detectable CO disk and dominates the S(0) emission; it contains 10 to 30% of the mass implied by CO observations (lower fractions will result if an ortho-para ratio below unity is adopted); concentrations of warmer gas (plausibly identified with active star forming regions) are ubiquitous in the inner disk and rarer in the outer disk and give rise to the S(1) emission. Given the density and temperature of the extended warm gas, this component can most likely be identified with extended low-density photon-dominated regions (PDRs) which form the warm envelopes of giant molecular clouds, heated by the local interstellar radiation field. Observations of Galactic molecular cloud edges in the 21 cm H I line reveal the presence of such warm envelopes by a “limb brightening” in H I (? (?; Van der Werf *et al.* 1988, 1989). Modeling of molecular cloud envelopes by ? (?) shows that the temperatures and densities estimated here are typical for such regions, especially in the zone where the transition from molecular to atomic hydrogen takes place. This interpretation of the present H_2 data in terms of extended diffuse PDRs in cloud envelopes is corroborated by the excellent numerical agreement with column density estimates based on [C II] 158 μm , the principal coolant for such regions. In addition, the fact that the implied ortho-para ratio is significantly lower than three also points to a low-density PDR origin of this emission, since shocks or hot, high-density PDRs would produce an ortho-para ratio of three (the high temperature thermal value) by spin exchange reactions with H and H^+ . In summary, the H_2 rotational lines in the disk of NGC 891 arise in warm, extended molecular cloud envelopes; these envelopes provide a physical link between the giant molecular clouds and the atomic medium in which they are embedded (?). The H_2 rotational lines provide unique diagnostics of these regions.

3. Vibrational H_2 emission in the nuclear region of the luminous merger NGC 6240

The vibrational H_2 emission of NGC 6240 has attracted attention because of its extraordinary luminosity: $7 \cdot 10^7 L_\odot$ in the $H_2 v = 1 \rightarrow 0$ S(1) line alone (for $H_0 = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and with no correction for extinction). This line contains 0.012% of the bolometric luminosity of NGC 6240, which is considerably higher than any other galaxy (?). Together the vibrational lines may account for 0.1% of the total bolometric luminosity.

Imaging of the $H_2 v = 1 \rightarrow 0$ S(1) emission from NGC 6240 has shown that the H_2 emission peaks *between* the two remnant nuclei of the merging system (?). This morphology provides a unique constraint on the excitation mechanism, since it argues against any scenario where the excitation is dominated by the stellar component (e.g., UV-pumping, excitation by shocks or X-rays from supernova remnants). Instead, the favoured excitation mechanism is slow shocks in the nuclear gas component, which, as shown by recent high resolution interferometry in the CO $J = 2 \rightarrow 1$ line (?), also peaks between the nuclei of NGC 6240. A multi-line H_2 vibrational spectrum (Fig. 2) confirms this excitation mechanism (?), in agreement with more limited spectroscopy by ? (?). The interpretation in terms of slow shocks also naturally accounts for the high ratio of H_2 line emission to infrared continuum emission, which is a characteristic of such shocks (e.g., ? (?).

What is the role of these shocks? In the shocks mechanical energy is dissipated and radiated away, mostly in spectral lines (principally H_2 , CO, H_2O and [O I] lines). This energy is radiated away at the expense of the orbital energy of the molecular clouds in the central potential well. Consequently, the dissipation of mechanical energy by the shocks will give rise to an infall of molecular gas to the centre of the potential well. Therefore, *the H_2 vibrational lines measure the rate of infall of molecular gas into the central potential well in NGC 6240.* This conclusion can be quantified by writing

$$L_{\text{rad}} = L_{\text{dis}}, \quad (3.1)$$

where L_{rad} is the total luminosity radiated by the shocks and L_{dis} the dissipation rate of mechanical energy, giving rise to a molecular gas infall rate \dot{M}_{H_2} given by

$$L_{\text{dis}} = \frac{1}{2} \dot{M}_{H_2} v^2, \quad (3.2)$$

where v is the circular orbital velocity at the position where the shock occurs.

Using a K -band extinction of 0^m15 (?), the total luminosity of H_2 vibrational lines from NGC 6240 becomes $7.2 \cdot 10^8 L_\odot$; inclusion of the purely rotational lines observed with the ISO SWS approximately doubles this number, so that $L_{\text{rad}} = 1.5 \cdot 10^9 L_\odot$.

In order to use this number to estimate \dot{M}_{H_2} , it is necessary to establish more accurately the fraction of the H_2 emission that is due to infalling gas. Observations with NICMOS on the Hubble Space Telescope (HST) by ? (?) provide the required information (Fig. 3). The NICMOS image shows that the emission consists of a number of tails (presumably related to the superwind also observed in $H\alpha$ emission), and concentrations associated with the two nuclei, and a further concentration approximately (but not precisely) between the two nuclei. The relative brightness of the H_2 emission from the southern nucleus is deceptive, since this nucleus is much better centred in the filter that was used for these observations than the other emission components, in particular the northern nucleus. Taking this effect into account, it is found that 32% of the total H_2 flux is associated with the southern nucleus, 16% is associated with the northern nucleus, and 12% with the component between the two nuclei, the remaining 40% being associated with extended emission. Using inclination-corrected circular velocities (from

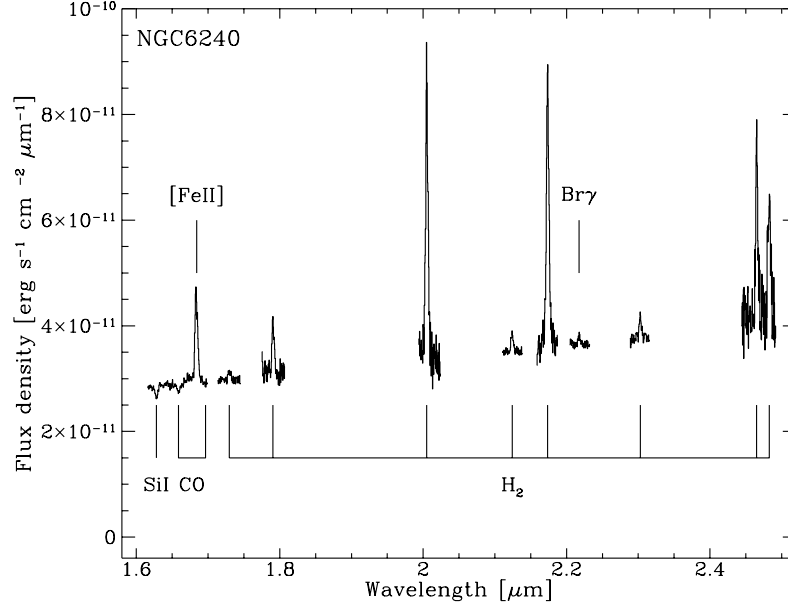


FIGURE 2. $H+K$ -band spectrum of NGC 6240 showing H_2 lines up to the $v = 1 \rightarrow 0$ S(9) line (?).

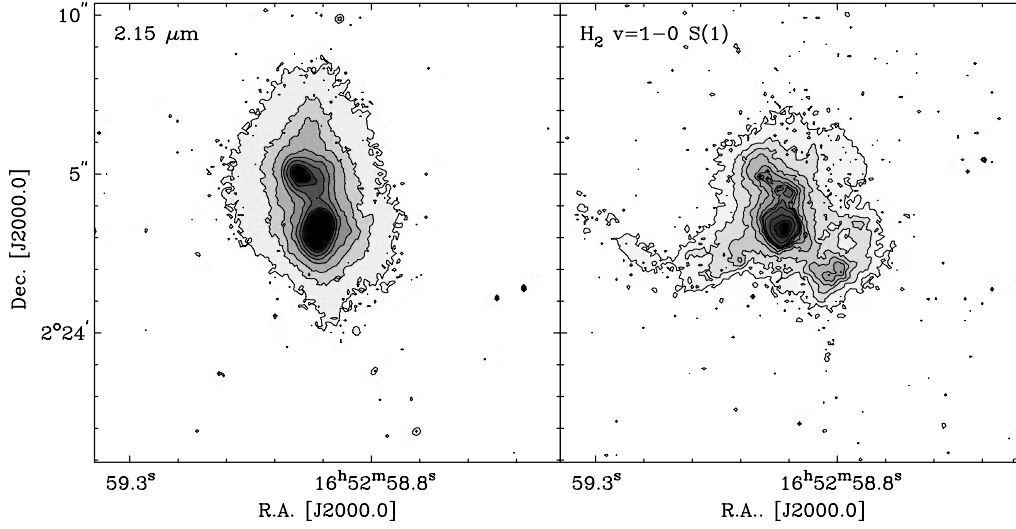


FIGURE 3. High resolution imaging of NGC 6240 in the H_2 $v = 1 \rightarrow 0$ S(1) line and the $2.15 \mu\text{m}$ continuum with NICMOS/HST (?).

? (?) of 270 and 360 km s^{-1} for the southern en northern nucleus respectively, and of 280 km s^{-1} for the central component (?), the mass infall rates derived using Eqs. (3.1) and (3.2) are $80 M_{\odot} \text{ yr}^{-1}$ for the southern nucleus, $22 M_{\odot} \text{ yr}^{-1}$ for the northern nucleus and $28 M_{\odot} \text{ yr}^{-1}$ for the central component.

The derived molecular gas inflow rate to the two nuclei is remarkably close to the mass consumption rate by star formation of approximately $60 M_{\odot} \text{ yr}^{-1}$, indicating that the H_2 emission from the nuclei directly measures the fueling of the starbursts in these regions.

The gas inflow towards the central component is more remarkable, since this component is not associated with a prominent stellar component. The gravitational potential at this position is therefore most likely dominated by the gas itself, which is also indicated by high resolution interferometry in the CO $J = 2 \rightarrow 1$ line (?). The absence of prominent star formation at this position shows that the central gas component is gravitationally stabilized, possibly by a high local shear. However, as the central gas column density increases, the shear must eventually be overcome and given the high gas density (and hence short free-fall time) an explosive starburst will result. In that stage NGC 6240 will become a true ultraluminous infrared galaxy.

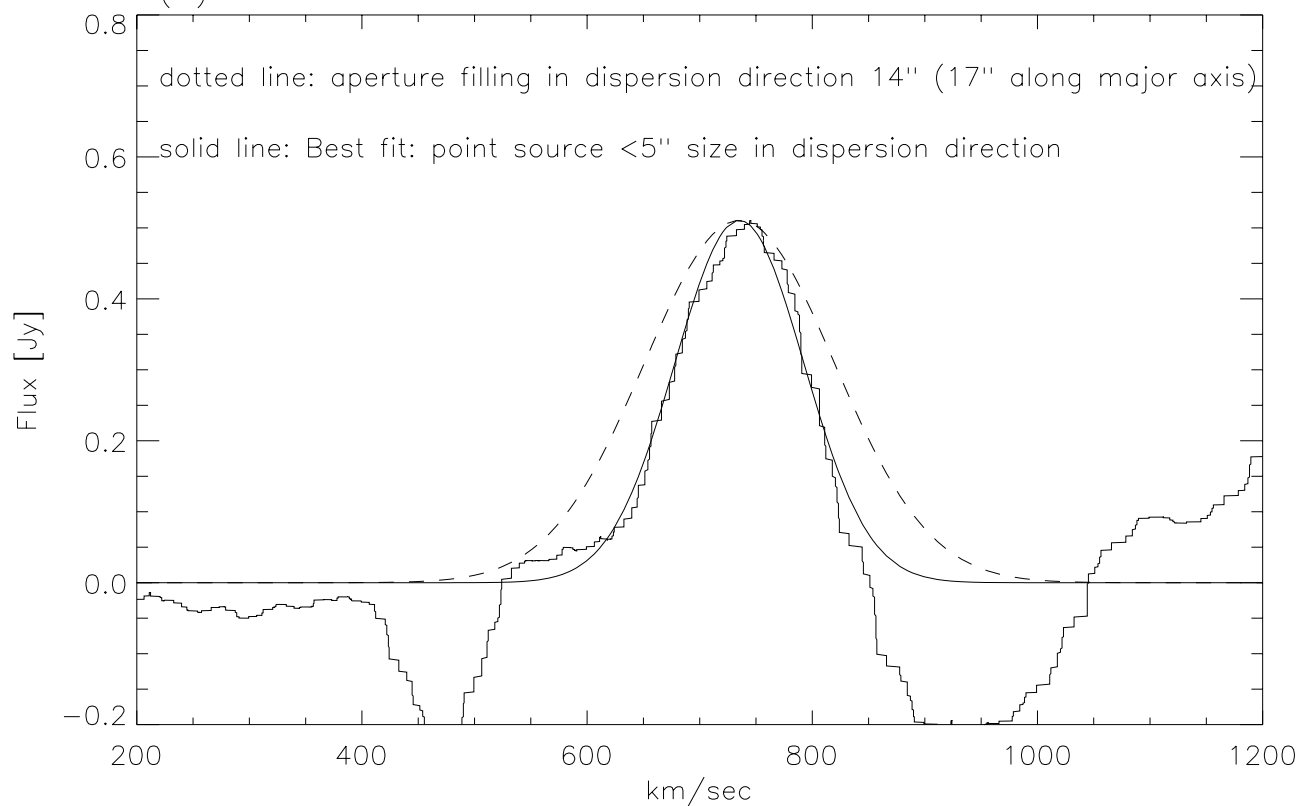
4. General implications

The two case studies discussed above illustrate the unique diagnostic power of H_2 emission lines provided that the excitation mechanism can be determined. The extended PDR emission detected in NGC 891 reveals the presence of a widespread fluorescent component, which is probably a common feature of star forming galaxies. For instance, ? (?) have detected extended diffuse H_2 vibrational emission from the inner Milky Way and argue that this emission is UV-excited; similarly, ? (?) argue for purely fluorescent extended H_2 emission in the moderate luminosity starburst galaxy NGC 253.

The configuration of shocked H_2 emission from a pronounced molecular gas concentration between the two nuclei as found in NGC 6240 is not unique either: it is also found in Arp 220 (?), and several other mergers. Perhaps the best case in point is the well-known merging “Antennae” system (NGC 4038 – 4039) where pronounced CO emission is found in the interaction zone between the two nuclei (?), which is the site of vigorous obscured star formation (?). These geometries may be due to the fact that the gas components are dissipative, and thus merge on a shorter time scale than the dissipationless stellar components, which merge by the slower process of dynamical friction. The H_2 emission thus provides unique insight into the physics of these gas-rich mergers.

I would like to thank my collaborators Frank Israel, Alan Moorwood, Tino Oliva, and Edwin Valentijn for discussions on this subject, and Guido Kusters for reducing the NICMOS data of NGC 6240.

H2 S(1) – NGC 891 OFF4– SPATIAL EXTENT in DISPERSION DIRECTION



H2 S(0) – NGC 891 OFF4– SPATIAL EXTENT in DISPERSION DIRECTION

